

BAB 1 PENDAHULUAN

1.1 Latar Belakang

Peningkatan polusi udara merupakan isu yang sangat penting di Indonesia. Menurut AQLI (2021) dalam Laporan Kualitas Udara Kehidupan, Indonesia termasuk dalam enam negara yang paling berkontribusi terhadap polusi udara global. Bersama dengan China, India, Pakistan, Bangladesh, dan Nigeria, Indonesia menyumbang 75% dari total beban polusi udara global karena tingkat polusi udara yang tinggi dan jumlah populasi yang besar. Dampak polusi udara ini dapat mengurangi angka harapan hidup hingga lebih dari enam tahun. Sementara, Laporan Kualitas Udara Dunia menunjukkan bahwa Indonesia menempati salah satu peringkat atas sebagai negara dengan udara paling tercemar. Indonesia juga berada di peringkat pertama sebagai negara dengan tingkat polusi udara tertinggi di Asia Tenggara (IQAir, 2022).

Polusi udara di Indonesia dapat berasal dari aktivitas manusia seperti transportasi, kegiatan industri, dan rumah tangga dan fenomena alam. Salah satu fenomena alam yang berkontribusi dalam peningkatan polusi udara di Indonesia ialah erupsi gunung api. Secara geografis, Indonesia terletak di kawasan *ring of fire* dunia dan memiliki 127 gunung api aktif yang menyebabkan jumlah kejadian erupsi gunung apinya termasuk tinggi jika dibandingkan negara lain. Hal ini mengakibatkan gunung api menjadi salah satu penyumbang besar dalam peningkatan polusi udara di Indonesia terutama untuk polutan yang bersumber dari erupsi seperti sulfur dioksida (SO_2), uap air (H_2O) dan karbon dioksida (CO_2). Salah satu zat polutan terbanyak yang berasal dari erupsi gunung api adalah SO_2 , gas ini merupakan gas berbahaya yang dapat menyebabkan gangguan pernafasan, sakit kepala dan dapat menyerang saraf manusia. Pada kadar yang melebihi ambang batas SO_2 dapat menyebabkan kematian. Selain berdampak pada manusia gas ini juga dapat mempengaruhi atmosfer. Dampak paling signifikan dari emisi vulkanik SO_2 ke atmosfer adalah oksidasi SO_2 menjadi asam sulfat (H_2SO_4), yang kemudian

dengan cepat mengembun di stratosfer dan membentuk aerosol sulfat cair yang dapat menyebabkan penipisan ozon (Bègue dkk., 2017). Selain itu, emisi dari gunung api yang mencapai stratosfer juga dapat memberikan efek pendinginan pada suhu global (Arfeuille dkk., 2013; Ravindra dkk., 2022; Ridley dkk., 2014). Dilihat dari potensi bahaya yang dapat ditimbulkan oleh konsentrasi SO_2 terhadap peningkatan polusi udara, tingginya jumlah kejadian erupsi gunung api di Indonesia dan sulitnya memprediksi polusi udara maka penelitian-penelitian terhadap prediksi polusi udara akibat erupsi gunung api sangat diperlukan.

Banyak cara yang dapat digunakan untuk memprediksi tingkat polusi udara di antaranya model statistik klasik, *machine learning*, *ensemble learning* hingga *deep learning*. Dalam penerapannya metode-metode ini saling berkaitan dalam satu proses yang disebut *data mining*. *Data mining* merupakan salah satu tahapan yang paling penting dari penemuan pengetahuan dalam proses basis data. Tujuan utama dari *data mining* adalah untuk mengekstraksi informasi yang berguna dari sejumlah besar data mentah dan mengubahnya menjadi bentuk yang dapat dimengerti untuk penggunaan yang efektif dan efisien (Archana dan Elangovan, 2014).

Penelitian menggunakan metode *data mining* dalam kajian prediksi polusi udara sudah pernah dilakukan di sebagian besar negara. Siwek dan Osowski (2016) melakukan penelitian menggunakan *data mining* untuk memprediksi polusi udara yang disebabkan oleh polutan atmosfer berupa nitrogen oksida (NO), sulfur dioksida (SO_2), ozon (O_3) dan materi partikulat $10\ \mu\text{m}$ (PM_{10}) di Warsaw, Polandia. Penelitian tersebut menyajikan dan membandingkan berbagai solusi dari sistem yang memprediksi polusi udara rata-rata harian dari PM_{10} , SO_2 , NO_2 , dan O_3 untuk hari berikutnya. Hasil dari penelitian tersebut berhasil mengkonfirmasi keunggulan akurasi prediksi dengan pendekatan menggunakan *data mining* untuk semua polutan yang dipertimbangkan. Namun, lokasi penelitian yang cukup kecil hanya pada satu kota menyebabkan terbatasnya data yang dapat digunakan untuk membangun model yang digunakan.

Selanjutnya, Yadav dkk. (2019) melakukan penelitian prediksi kualitas udara dengan menggunakan metode *data mining* berbasis *time series* di New Delhi,

India. Dalam penelitian digunakan *Apriori-based association rule mining algorithm* yang merupakan versi modifikasi dari *Continuous Target Sequential Pattern Discovery* (CTSPD) untuk menghasilkan seperangkat aturan asosiasi yang membantu memprediksi konsentrasi CO, O₃, NO₂, PM2.5 dan PM10. Dalam penelitian algoritma yang telah diusulkan juga dibandingkan dengan sistem prediksi yang ada yaitu *System of Air Quality and Weather Forecasting And Research* (SAFAR). Dari penelitian tersebut didapat kesimpulan bahwa algoritma yang diusulkan lebih akurat dalam memprediksi kualitas udara dari pada SAFAR. Namun, penelitian ini lebih berfokus menggunakan data yang didapat melalui sensor yang dioperasikan di permukaan daerah India saja sehingga dataset yang dapat digunakan termasuk terbatas.

Sementara itu, Huang dkk. (2021) menggunakan 5 model *data mining* berupa *long short-term memory* (LSTM), *backpropagation neural network* (BPNN), *adaptive network-based fuzzy inference system* (ANFIS), *generalized regression neural network* (GRNN), dan *support vector machines* (SVM) dengan bantuan algoritma *fuzzy c-means cluster* untuk menganalisis 13 seri *Air Quality Index* (AQI) standar, dan evaluasi komprehensif *fuzzy* digunakan untuk mengetahui polutan udara utama di 13 kota berbeda disekitar aglomerasi Beijing-Tianjin-Heibei. Zat polutan yang diukur mencakup enam polutan udara: PM2.5, PM10, CO, NO₂, SO₂ dan O₃. Penelitian tersebut berhasil menunjukkan bahwa sistem yang diusulkan dapat secara bersamaan memperoleh data yang efisien dan andal untuk pemantauan kualitas udara. Namun, penelitian tersebut masih menggabungkan sumber zat polutan antara polutan yang berasal dari kejadian bencana alam dan polutan yang berasal dari kegiatan antropogenik.

Metode *data mining* di Indonesia pernah dimanfaatkan dalam beberapa penelitian berkaitan dengan atmosfer contohnya dalam prediksi cuaca seperti yang dilakukan oleh Findawati dkk.(2019), prediksi polusi udara dari kegiatan antropogenik seperti yang dilakukan oleh Putra dan Sitanggang (2020) dan prediksi curah hujan yang dilakukan oleh Nanda dkk. (2022). Namun, hingga saat ini, belum ada penelitian di Indonesia yang secara khusus menerapkan teknik *data mining* untuk memprediksi

polusi udara akibat erupsi gunung api, dengan fokus pada satu jenis polutan seperti SO_2 . Hal ini menunjukkan adanya peluang penelitian baru yang dapat berkontribusi dalam pemantauan kualitas udara di wilayah rawan bencana erupsi gunung api seperti Indonesia.

Pada penelitian ini digunakan dua model *data mining* yaitu *Autoregressive Integrated Moving Average* (ARIMA) dan LSTM. Kedua model ini dipilih berdasarkan keunggulan masing-masing model dalam melakukan prediksi polusi udara dibandingkan model lain. Keunggulan model ARIMA terletak pada kemampuannya untuk mengatasi data non-stasioner dengan mengintegrasikan diferensiasi, yang sering kali diperlukan untuk dataset polusi udara yang menunjukkan tren dan pola musiman (Gourav et al., 2020). Selain itu, kemampuan model untuk memadukan komponen *autoregressive* (AR) dan *moving average* (MA) memungkinkan model ini untuk menangkap pola kompleks dalam data secara efektif (Gopu et al., 2021). Sementara model LSTM merupakan model yang dirancang untuk mengingat informasi dalam jangka waktu yang lama, yang sangat penting untuk menangkap *temporal dependency* atau ketergantungan terhadap waktu pada data polusi udara. Kemampuan ini memungkinkan LSTM untuk mengungguli model lain yang kesulitan mengatasi faktor ketergantungan terhadap waktu pada data jangka panjang, sehingga model ini dapat meningkatkan akurasi dalam prediksi jangka panjang (Baniasadi et al., 2023; Wang and Zhang, 2022). Pemanfaatan kedua metode ini diharapkan dapat menghasilkan model prediksi yang tidak hanya akurat tetapi juga adaptif terhadap dinamika data polusi udara akibat erupsi gunung api.

1.2 Tujuan Penelitian

Tujuan penelitian ini adalah menganalisis persebaran konsentrasi SO_2 akibat erupsi gunung api di Indonesia, memprediksi tingkat konsentrasi SO_2 tersebut menggunakan model *data mining* ARIMA dan LSTM serta mengukur tingkat keakuratan dari model yang digunakan dalam memprediksi tingkat konsentrasi SO_2 akibat erupsi gunung api di Indonesia.

1.3 Manfaat Penelitian

Penelitian ini memberikan kontribusi dalam pengembangan bidang ilmu pengetahuan berkaitan dengan *data mining* terutama untuk model ARIMA dan LSTM, dan analisis persebaran polusi udara konsentrasi SO₂ di wilayah rawan bencana erupsi gunung api, menjadi acuan bagi pemerintah dalam pengambilan keputusan dalam mengantisipasi dampak buruk dari erupsi gunung api terhadap kualitas udara di Indonesia dan dapat menjadi acuan untuk penelitian-penelitian berikutnya.

1.4 Ruang Lingkup dan Batasan Penelitian

Penelitian ini dibatasi untuk wilayah keseluruhan negara Indonesia. Data yang digunakan adalah data harian *total column amount* SO₂ dari satelit OMI-AURA NASA dengan kurun waktu 18 tahun (2005-2022). Penelitian ini menggunakan model ARIMA dan LSTM sebagai teknik membuat prediksi dengan batas prediksi 5 tahun (2023-2027).

